Homework 4 ***DRAFT*** **EECS 651** **Feb. 10, 2001**

1. *In this problem, you will estimate the SNR of an audio compact disc (CD) and compare the result to that attainable by if other quantization methods were used. Here's what you need to know. Audio CD's are encoded with a fixed-rate, uniform scalar quantizer (USQ) with rate 16. Let [-L,L] denote the support region of the uniform scalar quantizer. Assume the audio source is (first-order) autogressive, stationary and Gaussian with zero mean, variance σ^2, and correlation coefficient ρ=0.9.*

(a) *Assuming $L = 10\sigma$ (I'm guessing this is reasonable), estimate the SNR in dB.*

In this case, the USQ step size is $\Delta = \frac{2L}{2^{16}} = \frac{20\,\sigma}{2^{16}}$. The MSE is $D \cong \frac{\Delta^2}{12} = \frac{20^2\,\sigma^2}{12\times 2^{32}}$,

and the SNR in dB is $10 \log_{10} \frac{\sigma^2}{D} \cong 10 \log_{10} \frac{12\times 2^{32}}{20^2} = 10 \log_{10} 1.28 \times 10^6 = \mathbf{81.1\ dB}$

(b) *Estimate the SNR in dB assuming (for this part only) that the audio source is uniformly distributed between -L and L.*

In this case the largest SNR with a rate R, fixed-rate USQ is $S_{usq}(R) = 6.02$ R. So the answer is

$S_{usq}(16) = \mathbf{96.3\ dB}$

(c) *Estimate the largest SNR in dB that could be attained by any fixed-rate scalar quantizer (uniform or nonuniform) with rate 16.*

From Zador's theorem the largest SNR attainable with scalar quantization with rate 16 is, approximately,

$$\mathcal{Z}(1,16) = 6.02 \times 16 - 10 \log_{10} m_1^* \beta_1$$

We know that $m_1^* = \frac{1}{12}$ and for the Gaussian source $\beta_1 = 2\pi\, 3^{3/2} = 32.6$. Substituting these gives

92.0 dB

(d) Estimate the largest SNR in dB that could be attained by any fixed-rate k-dimensional vector quantizer with rate 16, for k = 2,3,4,8.

From Zador's theorem the largest SNR attainable with k-dimensional VQ with rate 16 is, approximately,

$$\mathcal{Z}(k,16) = 6.02 \times 16 - 10 \log_{10} m_k^* \beta_k$$

For the AR Gaussian source we know $\beta_k = 2\pi \left(\frac{k+2}{k}\right)^{(k+2)/2} (1-\rho^2)^{(k-1)/k}$

Except for k = 2, we don't know m_k^* exactly. However, we can use one of the approximate values given in the lecture notes. I'll use the "conjectured lower bound", but it would be OK to use the "best known" or the NMI of a sphere.

****** This gives the following table, the last column of which gives the answers to this part:

k	m_k^*	β_k	$m_k^* \beta_k$	$10 \log_{10} m_k^*\beta_k$	$\mathcal{Z}(k,16)$
1	.0833	32.6	2.72	4.35	92.0 dB
2	.0802	10.95	.88	-.56	**96.9 dB**
3	.0785	7.45	.58	-2.32	**98.6 dB**
4	.0761	6.10	.46	-3.33	**99.7 dB**
8	.0716	4.48	.32	-4.94	**101.3 dB**
∞	.0585	3.25	.19	-7.21	103.5 dB

The first and last rows are not required. They are just there for comparison and completeness.

(e) *Repeat (d) assuming the vq can have any dimension whatsover.*

As given in the table above, the largest SNR with VQ of any dimension and rate 16 is **103.5 dB**

(f) *Make a table showing how much more (in percent) audio could be stored on a CD with the methods of parts of (c), (d) and (e) than the with uniform scalar quantization with $L = 10\ \sigma$, assuming these other methods attain the same SNR as in Part (a).*

We know that SNR increases 6.02 dB per one bit increase of the rate. The method of (c) (nonuniform scalar quantization) gains 92.0-81.1 = 8.9 dB. Equivalently, if we keep the SNR the same, the method of (c) permits a reduction in rate of 8.9/6.02 = 1.48 bits/sample. This is a 1.48/16 = 9.3% decrease in rate, which enables a 9.3% increase in the stored amount of music. The following table gives the increases for all three cases

k	dB gain	rate reduction	%rate reduction=%increase in storage
1	8.9 dB	1.48 b/sample	**9.2%**
2	15.8 dB	2.62 b/sample	**16.4%**
3	17.5 dB	2.91 b/sample	**18.2%**
4	18.6 dB	3.09 b/sample	**19.3%**
8	20.2 dB	3.56 b/sample	**21.0%**
∞	22.4 dB	3.72 b/sample	**23.2%**

2. *Derive the formula for the Zador factor β_k assuming $\underline{X}$ is a k-dimensional IID random vector with Laplacian marginal densities.*

Consider the Laplacian density $f_X(x) = \frac{1}{\sqrt{2}}\, e^{-\sqrt{2}|x|}$, with variance 1. (Since β_k does not depend on the variance we need only consider the case of variance 1.) Since $\underline{X}$ is IID, we can use Property 5 of the lecture notes, which shows

$$\beta_k = \frac{1}{\sigma^2}\left(\int f_1(x)^{k/k+2}\,dx\right)^{k+2} = \left(\int_{-\infty}^{\infty} 2^{-k/2(k+2)}\, e^{-\sqrt{2}|x|k/(k+2)}\,dx\right)^{k+2}$$

$$= \left(2^{-k/2(k+2)} \int_{-\infty}^{\infty} \frac{k}{(k+2)\sqrt{2}}\, e^{-\sqrt{2}|x|k/(k+2)}\,dx\ \frac{(k+2)\sqrt{2}}{k}\right)^{k+2}$$ we've massaged this so only a Laplacian density lies within the integral

$$= \left(2^{-k/2(k+2)}\ \frac{(k+2)\sqrt{2}}{k}\right)^{k+2}$$ since the Laplacian density integrates to 1

$$= 2^{-k/2+(k+2)/2}\left(\frac{k+2}{k}\right)^{k+2} = 2\left(\frac{k+2}{k}\right)^{k+2}$$

3. *A VQ is needed for a (first-order) autoregressive, stationary Gaussian source with correlation coefficient ρ = .95. It must have rate 4 or less and signal-to-noise ratio 32.5 dB or more. Determine whether or not there exists a suitable VQ. If yes, estimate the smallest possible dimension.*

From the SNR version of Zador's formula, the largest SNR achievable by a VQ with any dimension and rate R (assuming R is large) is

$$S(R) \cong 6.02\, R - 10 \log_{10} m^*_\infty \beta_\infty .$$

R = 4 is large enough for the formula to apply. We know $m^*_\infty = 1/(2\pi e)$, and for a Gauss-Markov source $\beta_\infty = 2\pi e(1-\rho^2)$, where ρ = correlation coefficient. Hence,

$$S(4) \cong 34.2 \text{ dB} .$$

Since this is greater than 32.5, **there is a suitable VQ** with rate 4.

To estimate the smallest possible dimension, we use Zador's formula (for SNR) for the largest SNR achievable by a VQ with dimension k and rate R:

$$S(k,R) \cong 6.02\, R - 10 \log_{10} m^*_k \beta_k$$

where for this source $\beta_k = 2\pi\left(\frac{k+2}{k}\right)^{(k+2)/2}(1-\rho^2)^{(k-1)/k}$. Since R = 4, is large enough, we merely need to find the smallest k such that $S(k,4) \geq 32.5$ dB. As we try to do so, we realize that we only know m^*_k for k = 1, 2. However, upper and lower bounds to m^*_k were given in the notes. Using these we make the following table:

k	beta	conj'd lower bound to m^*_k	upper bound to m^*_k	lower bound to $\delta(k,R)$	upper bound to $\delta(k,R)$	upper bound to S(k,R)	lower bound to S(k,R)
1	32.65	0.0833	0.0833	1.06E-02	1.06E-02	19.74	19.74
2	7.85	0.0802	0.0802	2.46E-03	2.46E-03	26.09	26.09
3	4.77	0.077875	0.0785	1.45E-03	1.46E-03	28.38	28.35
4	3.70	0.0761	0.0766	1.10E-03	1.11E-03	29.59	29.56
5	3.17	0.0747	0.0756	9.25E-04	9.36E-04	30.34	30.29
6	2.85	0.0735	0.0742	8.19E-04	8.27E-04	30.87	30.82
7	2.65	0.0725	0.0731	7.50E-04	7.56E-04	31.25	31.22
8	2.50	0.0716	0.0717	6.99E-04	7.00E-04	31.55	31.55
12	2.19	0.0692	0.0701	5.91E-04	5.99E-04	32.28	32.22
13	*2.14*	*0.0676*	*0.0701*	*5.66E-04*	*5.87E-04*	*32.47*	*32.31*
14	*2.11*	*0.0676*	*0.0701*	*5.56E-04*	*5.77E-04*	*32.55*	*32.39*
15	*2.07*	*0.0676*	*0.0701*	*5.47E-04*	*5.68E-04*	*32.62*	*32.46*
16	2.05	0.0676	0.0683	5.40E-04	5.46E-04	32.68	32.63
24	1.91	0.0656	0.0658	4.90E-04	4.91E-04	33.10	33.09

We weren't given bounds to m^*_k for k = 13, 14, 15 so for the lower bound I used the value for k=16 and for the upper bound I used the value for k = 12.

We see from that k = 13 won't work (won't give SNR ≥ 32.5), k = 14 and 15 might work, and k = 16 definitely works. So our answer is that **the least value of k is 14, 15 or 16**.

4. *Show that* $\beta_1 \geq \frac{1}{\sigma^2} 2^{2h}$ *where* $h = -\int_{-\infty}^{\infty} f(x) \log_2 f(x)\, dx$.

(Hints: You might try Jensen's inequality or $\ln x \leq x-1$. *Also,* $h = E[\frac{3}{2} \log f^{-2/3}(X)]$.*)*

Derivation based on Jensen's inequality: Jensen's inequality says that if a function g is convex $\cap$, then $E\, g(X) \leq g(EX)$.

$$h = -\int_{-\infty}^{\infty} f(x) \log_2 f(x)\, dx = \frac{3}{2} E[\log f^{-2/3}(X)] = \frac{3}{2} E[\log Y)], \quad \text{where } Y = f^{-2/3}(X)$$

$$\leq \frac{3}{2} \log E\, Y = \frac{3}{2} \log E[f^{-2/3}(X)] \quad \text{from Jensen's inequaltiy and fact that log is convex} \cap$$

$$= \frac{3}{2} \log \int_{-\infty}^{\infty} f^{-2/3}(x) f(x)\, dx = \frac{3}{2} \log \int_{-\infty}^{\infty} f^{1/3}(x)\, dx = \frac{1}{3} \log \left(\int_{-\infty}^{\infty} f^{1/3}(x)\, dx \right)^3 = \frac{1}{2} \log \sigma^2 \beta_1$$

$$\Rightarrow \quad \beta_1 = \frac{1}{\sigma^2} 2^{2h}$$

Derivation based on $\ln x \leq x-1$:

$$h = -\int_{-\infty}^{\infty} f(x) \log_2 f(x)\, dx = \frac{3}{2} E[\log_2 f^{-2/3}(X)] = \frac{3}{2} E[\ln f^{-2/3}(X)] \frac{1}{\ln 2}$$

$$= \frac{3}{2} E \ln \frac{f^{-2/3}(X)}{(\sigma^2\beta_1)^{1/3}} \frac{1}{\ln 2} + \frac{3}{2} \ln (\sigma^2\beta_1)^{1/3} \frac{1}{\ln 2}$$

$$\leq \frac{3}{2} E\left(\frac{f^{-2/3}(X)}{(\sigma^2\beta_1)^{1/3}} - 1 \right) \frac{1}{\ln 2} + \frac{1}{2} \ln (\sigma^2\beta_1) \frac{1}{\ln 2}, \quad \text{using } \ln x \leq x-1$$

$$= \frac{1}{2} \ln (\sigma^2\beta_1) \frac{1}{\ln 2} \quad \text{because } E f^{-2/3}(X) = \int_{-\infty}^{\infty} f^{-2/3}(x) f(x)\, dx = (\sigma^2\beta_1)^{1/3}$$

$$= \frac{1}{2} \log \sigma^2\beta_1$$

$$\Rightarrow \quad \beta_1 = \frac{1}{\sigma^2} 2^{2h}$$

5. *Do there exist prefix codes with the following sets of codeword lengths?*

(a) {2,2,3,3,3,5,6,6,6,6,7} .

The Kraft inequality is less than 1, so there exists a prefix code with these lengths.

(b) {2,3,3,3,4,4,4,4,4}

The Kraft inequality is less than 1, so there exists a prefix code with these lengths.

(c) {2,2,2,3,3}

The Kraft inequality equals 1, so there exists a prefix code with these lengths.

(d) *For any set for which there does exist a code, draw the binary tree of a code with these lengths.*

There are many possibilities.

(a) {00, 01, 100, 101, 11000,110010, 110011, 1111111}

(b) {00, 100, 101, 110, 1110, 1111, 0100, 0101, 0110}

(c) {00, 01, 10, 110, 111}

6. *Consider an IID source with the following set of probabilities:*

$\{.25,.2,.1,.1,.1,.1,.05,.05,.05\}$.

(a) *Find the entropy of the source.* **H = 2.94**

(b) *Find two different prefix codes (first-order) with minimum rate. The codes should have different sets of lengths.*

Using Huffman's algorithm and breaking ties in different ways, we find the following two codes with minimum average length: (these are not the only possibilities)

code 1:	{ 00, 10, 110, 111, 0100, 0101, 0110, 01110, 01111 }
lengths	{ 2, 2, 3, 3, 4, 4, 4, 5, 5 }
code 2:	{ 00, 010, 100, 101, 110, 111, 0110, 01110, 01111 }
lengths	{ 2, 3, 3, 3, 3, 3, 4, 5, 5 }

Both codes have rate = 3.

(c) *Compare the entropy and the rate of the codes found in (b). Do they differ by a "reasonable" amount?*

The rate of the code is .06 larger than the entropy. We know that it could not be less than the entropy, and we know that it can't be more than the entropy plus 1. Moreover, we know that it can't be larger than the entropy by more than the largest probability, which is .25. So being larger than entropy by .06 is quite reasonable.

7. *Find an example of a source for which $R^* \geq H + .9$. Hint: a binary source will suffice. This shows that R^* can be very close to $H + 1$.*

Let X be a binary IID source with $P_0 = .99$, $P_1 = .01$. Then $H = .081$, $R^* = 1$.

So $1 = R^* \geq H + .9 = .981$

8. *Show by example that a prefix code with lengths $l_i = \lceil -\log_2 P_i \rceil$ does not necessarily have minimum average length.*

There are many possible examples. Here's one. Let $M = 2$, $P_1=7/8$, $P_2 = 1/8$. Then

$\lceil -\log_2 P_1 \rceil = 1$ and $\lceil -\log_2 P_i \rceil = 3$.

But in an optimum code, both codewords would have length 1. So choosing $l_i = \lceil -\log_2 P_i \rceil$ does not yield an optimum prefix code.

9. *Show that if each probability in the set $\{P_1,...,P_M\}$ is a negative power of 2, then the Shannon code is an optimal prefix code.*

****** Suppose $P_i = 2^{n_i}$ for each i. Then the average length of the Shannon code is

$$\sum_{i=1}^{M} P_i l_i = \sum_{i=1}^{M} P_i \lceil -\log_2 P_i \rceil = \sum_{i=1}^{M} P_i \lceil -n_i \rceil = \sum_{i=1}^{M} P_i (-n_i) = -\sum_{i=1}^{M} P_i \log_2 P_i = H$$

****** Since the average length of the Shannon code is H and since no code can have average length less than H, the Shannon code is optimal.